A MODIFIED SEMI–IMPLICIT EULER-MARUYAMA SCHEME FOR FINITE ELEMENT DISCRETIZATION OF SPDES WITH ADDITIVE NOISE

GABRIEL J LORD [†,‡] AND ANTOINE TAMBUE [†,§,*]

Abstract. We consider the numerical approximation of a general second order semi–linear parabolic stochastic partial differential equation (SPDE) driven by additive space-time noise. We introduce a new scheme using in time a linear functional of the noise with a semi–implicit Euler–Maruyama method and in space we analyse a finite element method although extension to finite differences or finite volumes would be possible. We give the convergence proofs in the root mean square L^2 norm for a diffusion reaction equation and in root mean square H^1 norm in the presence of advection. We examine the regularity of the initial data, the regularity of the noise and errors from projecting the noise. We present numerical results for a linear reaction diffusion equation in two dimensions as well as a nonlinear example of two-dimensional stochastic advection diffusion reaction equation. We see from both the analysis and numerics that we have better convergence properties over the standard semi–implicit Euler–Maruyama method.

Key words. Parabolic stochastic partial differential equation, Finite element, Modified Semi–implicit Euler–Maruyama, Strong numerical approximation, Additive noise.

1. Introduction.

We analyse the strong numerical approximation of Ito stochastic partial differential equation defined in $\Omega \subset \mathbb{R}^d$. Boundary conditions on the domain Ω are typically Neumann, Dirichlet or some mixed conditions. We consider

$$dX = (AX + F(X))dt + dW, \qquad X(0) = X_0, \qquad t \in [0,T], \quad T > 0 \tag{1.1}$$

in a Hilbert space $H = L^2(\Omega)$. Here A is the generator of an analytic semigroup $S(t) := e^{tA}, t \geq 0$ with eigenfunctions e_i and eigenvalues λ_i, $i \in \mathbb{N}^d$. F is a nonlinear function of X and possibly ∇X and the noise term $W(x,t)$ is a Q-Wiener process that is white in time and defined on a filtered probability space $(\mathbb{D}, \mathcal{F}, \mathbf{P}, \{F_t\}_{t\geq 0})$. The noise can be represented as

$$W(x,t) = \sum_{i \in \mathbb{N}^d} \sqrt{q_i} e_i(x) \beta_i(t), \tag{1.2}$$

where $q_i \geq 0$, $i \in \mathbb{N}^d$ and β_i are independent and identically distributed standard Brownian motions. Precise assumptions on A, F and W are given in Section 2 and under these type of technical assumptions it is well known, see [5, 18, 4] that the unique mild solution is given by

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))ds + O(t) \tag{1.3}$$

with the stochastic process O given by the stochastic convolution

$$O(t) = \int_0^t S(t-s)dW(s). \tag{1.4}$$

[†]Department of Mathematics and Maxwell Institute, Heriot-Watt University, Edinburgh EH14 4AS, UK

[‡]`g.j.lord@hw.ac.uk`

[§]`at150@hw.ac.uk (Antoine.Tambue@math.uib.no)`. Corresponding author.

[*]Department of Mathematics, University of Bergen, Johannes Bruns Gate 12, 5020 Bergen ,Norway.

Typical examples of the above type of equation are stochastic (advection) reaction diffusion equations where $A = \Delta$ arising from example in pattern formation in physics and mathematical biology. We illustrate our work with both a simple reaction diffusion equation where we can construct an exact solution

$$dX = (D\Delta X - \lambda X)\,dt + dW \tag{1.5}$$

as well as the stochastic advection reaction diffusion equation

$$dX = \left(D\Delta X - \nabla\cdot(\mathbf{q}X) - \frac{X}{|X|+1}\right)dt + dW \tag{1.6}$$

where $D > 0$ is the diffusion coefficient, $\mathbf{q}$ is the Darcy velocity field [3] and λ is a constant depending of the reaction.
The study of numerical solutions of SPDEs is an active area of research and there is a growing literature on numerical methods for SPDEs. We analyse convergence of a finite element discretization in space combined with a semi-implicit discretization in time that uses a linear functional to approximate the noise. This approach extends the analysis of Jentzen [11, 10, 12] which is based on a Fourier spectral discretization. This discretization diagonalizes the linear operator A and then exploits that in each Fourier mode the noise is an Ornstein–Uhlenbeck process and hence the variance is known. For complex domains or mixed boundary conditions a Fourier spectral approach is not feasible and, for example, finite element (finite difference or finite volume) discretization is preferred even they are less accurate compared to spectral methods, however this destroys the diagonalization of the linear operator.
Note that even if the eigenfunctions of the linear operator are known explicitly for the given domain and boundary conditions it may be that a finite element (or finite volume) discretization is prefered over a Fourier spectral method. For example where the solution has steep fronts finite elements and finite volumes are efficient and there is no issue with any Gibbs phenomena as in Fourier spectral approaches. Such solutions arise, for example, in heterogeneous porous media. In this case the Darcy velocity $\mathbf{q}$ in (1.6) is solved for from the following system

$$\begin{cases} \nabla\cdot\mathbf{q} = 0 \\ \mathbf{q} = -\dfrac{\mathbf{k}(x)}{\mu}\nabla p, \end{cases} \tag{1.7}$$

where $\mathbf{k}$ is the permeability tensor, p is the pressure and μ is the dynamic viscosity of the fluid [2]. Typically this would be solved using finite elements or finite volumes. Furthermore, due to the heterogenous nature of the permeability such problems often naturally give rise to non-uniform grids. Our approach, with a projection of the noise onto a standard finite element grid, allows practitioners to simply adapt their existing codes to examine the effects of stochastic forcing.
We perform our analysis for the case of finite elements and numerically we also examine finite volumes. Our work differs from other finite element discretizations [1, 14, 15] where the approach to the noise is to consider it directly in the finite element space. We follow more closely [22, 23, 8] and introduce a projection onto a finite number modes and a projection onto the finite element space. The aim is to gain the flexibility of the finite element (finite volume) discretization to deal with flow and transport problems (1.6)- (1.7), complex boundaries, mixed boundary conditions and inhomogeneous boundary conditions as well as the overcoming the order barrier as in [12].

We also show in Section 4.2 that equipped with the eigenvalues and eigenfunctions of the operator Δ with Neumann or Dirichlet boundary conditions, we can apply the new scheme with mixed boundary conditions for the operator $A = D\Delta$ without explicitly having the eigenvalues and eigenfunctions of A. In general for an operator A with given eigenvalues and eigenfunctions for Neumann or Dirichlet boundary conditions, the new scheme can be used for the operator A with mixed bounded conditions without explicit eigenvalues and eigenfunctions.
We give convergence proofs in root mean square $L^2(\Omega)$ norm for reaction-diffusion equations and in root mean square $H^1(\Omega)$ norm for advection reaction-diffusion for spatially regular noise. The smoothing effect of the semigroup generated by the operator A in the SPDE (1.1) and various semigroup estimates play an important role in the proofs.
The paper is organised as follows. In Section 2 we present the numerical scheme and assumptions that we make on the linear operator, nonlinearity and the noise. We then state and discuss our main results. In Section 3.2 and Section 3.3 we present the proof of the convergence theorems. We end by presenting some simulations in Section 4, these are applied both to a linear example where we can compute an exact solution as well as a more realistic model coming from model of the advection and diffusion of a solute in a porous media with a non-linear reaction term.

2. Numerical scheme and main results. Let us start by presenting briefly the notation for the main function spaces and norms that we use in the paper. We denote by $\|\cdot\|$ the norm associated to the inner product $(\cdot,\cdot)$ of the $\mathbb{R}-$Hilbert space $H = L^2(\Omega)$. For a Banach space $\mathcal{V}$ we denote by $\|\cdot\|_{\mathcal{V}}$ the norm of the space $\mathcal{V}$, $L(\mathcal{V})$ the set of bounded linear mapping from $\mathcal{V}$ to $\mathcal{V}$, $L^{(2)}(\mathcal{V})$ the set of bounded bilinear mapping from $\mathcal{V}\times\mathcal{V}$ to $\mathbb{R}$ and $L_2(\mathbb{D},\mathcal{V})$ the space defined by

$$L_2(\mathbb{D},\mathcal{V}) = \left\{ v \text{ random variable with value in } \mathcal{V}:\ \mathbf{E}\|v\|^2_{\mathcal{V}} = \int_{\mathbb{D}} \|v(\omega)\|^2_{\mathcal{V}} d\mathbf{P}(\omega) < \infty \right\}.$$

Throughout the paper we assume that Ω is bounded and has a smooth boundary or is a convex polygon. For convenience of presentation we take A to be a self adjoint second order operator as this simplifies the convergence proof. More precisely

$$A = \nabla\cdot\mathbf{D}\nabla(.) + D_{0,0}\mathbf{I} = \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(D_{i,j}\frac{\partial}{\partial x_j}\right) + D_{0,0}\mathbf{I}, \tag{2.1}$$

where we assume that $D_{i,j} \in L^{\infty}(\Omega)$, and that there exists a positive constant $c_1 > 0$ such that

$$\sum_{i,j=1}^{d} D_{i,j}(x)\xi_i\xi_j \geq c_1|\xi|^2 \qquad \forall \xi \in \mathbb{R}^d \quad x \in \overline{\Omega} \quad c_1 > 0. \tag{2.2}$$

Under condition (2.2), it is well known (see [7]) that the linear operator A generates an analytic semigroup $S(t) \equiv e^{tA}$. We introduce two spaces $\mathbb{H}$ and V where $\mathbb{H} \subset V$, that depend on the choice of the boundary conditions for the domain of the operator A and for the variational form associated to the operator A. For Dirichlet boundary conditions we let

$$V = \mathbb{H} = H^1_0(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega\}.$$

For Robin boundary conditions (Neumann boundary condition being a particular case) $V = H^1(\Omega)$ and

$$\mathbb{H} = \left\{ v \in H^1(\Omega) : \partial v/\partial \nu_A + \sigma v = 0 \quad \text{on} \quad \partial\Omega \right\}, \qquad \sigma \in \mathbb{R}.$$

Functions in $\mathbb{H}$ can satisfy the boundary conditions and with $\mathbb{H}$ in hand we can characterize the domain of the operator $(-A)^{r/2}$ and have the following norm equivalence [16, 6] for $r \in \{1,2\}$

$$\|v\|_{H^r(\Omega)} \equiv \|(-A)^{r/2} v\| =: \|v\|_r \qquad \forall \quad v \in \mathcal{D}((-A)^{r/2}) = \mathbb{H} \cap H^r(\Omega).$$

In the Banach space $\mathcal{D}((-A)^{\alpha/2})$, $\alpha \in \mathbb{R}$, we use the notation $\|.\|_\alpha := \|(-A)^{\alpha/2}.\|$. We recall some basic properties of the semi group $S(t)$ generated by A.

Proposition 2.1. ***[Smoothing properties of the semi group[9]]*** *Let $\alpha > 0$, $\beta \geq 0$ and $0 \leq \gamma \leq 1$, then there exist $C > 0$ such that*

$$\begin{aligned} \|(-A)^{\beta} S(t)\|_{L(L^2(\Omega))} &\leq C t^{-\beta} \quad \textit{for} \quad t > 0 \\ \|(-A)^{-\gamma}(I - S(t))\|_{L(L^2(\Omega))} &\leq C t^{\gamma} \quad \textit{for} \quad t \geq 0. \end{aligned}$$

In addition,

$$\begin{gathered} (-A)^{\beta} S(t) = S(t)(-A)^{\beta} \quad \textit{on} \quad \mathcal{D}((-A)^{\beta}) \\ \textit{If} \quad \beta \geq \gamma \quad \textit{then} \quad \mathcal{D}((-A)^{\beta}) \subset \mathcal{D}((-A)^{\gamma}), \\ \|D_t^l S(t) v\|_\beta \leq C t^{-l-(\beta-\alpha)/2} \|v\|_\alpha, \;\; t > 0, \; v \in \mathcal{D}((-A)^{\alpha/2}) \;\; l = 0, 1, \end{gathered}$$

where $D_t^l := \dfrac{d^l}{dt^l}$.

We consider discretization of the spatial domain by a finite element triangulation. Let $\mathcal{T}_h$ be a set of disjoint intervals of Ω (for $d=1$), a triangulation of Ω (for $d=2$) or a set of tetrahedra (for $d=3$). Let $V_h \subset V$ denote the space of continuous functions that are piecewise linear over the triangulation $\mathcal{T}_h$. To discretize in space we introduce two projections. Our first projection operator P_h is the $L^2(\Omega)$ projection onto V_h defined for $u \in L^2(\Omega)$ by

$$(P_h u, \chi) = (u, \chi) \qquad \forall \; \chi \in V_h. \tag{2.3}$$

Then $A_h : V_h \to V_h$ is the discrete analogue of A defined by

$$(A_h \varphi, \chi) = (A\varphi, \chi) \qquad \varphi, \chi \in V_h. \tag{2.4}$$

We denote by S_h the semigroup generated by the operator A_h.

The second projection P_N, $N \in \mathbb{N}$ is the projection onto a finite number of spectral modes e_i defined for $u \in L^2(\Omega)$ by

$$P_N u = \sum_{i \in \mathcal{I}_N} (e_i, u) e_i,$$

where $\mathcal{I}_N = \{1, 2, ..., N\}^d$. The semi–discrete in space version of the problem (1.1) is to find the process $X^h(t) = X^h(., t) \in V_h$ such that for $t \in [0, T]$,

$$dX^h = (A_h X^h + P_h F(X^h))dt + P_h P_N dW, \qquad X^h(0) = P_h X_0. \tag{2.5}$$

We denote by $\overline{X}^h$ the solution of the random system

$$\overline{X}^h(t) = S_h(t)X^h(0) + \int_0^t S_h(t-s)F(X^h(s))ds.$$

As in (1.3), by splitting we have

$$X^h(t) = \overline{X}^h(t) + P_h P_N O(t).$$

We now discretize in time by a semi–implicit method to get the fully discrete approximation of $\overline{X}^h$ defined by Z_m^h

$$Z_m^h = S_{h,\Delta t}^m P_h X_0 + \Delta\, t \sum_{k=0}^{m-1} S_{h,\Delta t}^{(m-k)} P_h F(Z_k^h + P_h P_N O(t_k)). \tag{2.6}$$

where

$$S_{h,\Delta t} := (\mathrm{I} - \Delta t\, A_h)^{-1}. \tag{2.7}$$

It is straightforward to show that

$$Z_{m+1}^h = S_{h,\Delta t}\left(Z_m^h + \Delta t P_h F(Z_m^h + P_h P_N O(t_m))\right). \tag{2.8}$$

Finally we can define our approximation X_m^h to $X(t_m)$, the solution of equation (1.1) by

$$X_m^h = Z_m^h + P_h P_N O(t_m). \tag{2.9}$$

Therefore

$$X_{m+1}^h = S_{h,\Delta t}\left(X_m^h + \Delta t\, P_h F(X_m^h) - P_h P_N O(t_m)\right) + P_h P_N O(t_{m+1}), \tag{2.10}$$

where according to (1.4), we generate $O(t_{m+1})$ from $O(t_m)$ by

$$O(t_{m+1}) = e^{\Delta t A} O(t_m) + \int_{t_m}^{t_{m+1}} e^{(t_{m+1}-s)A} dW(s). \tag{2.11}$$

If we assume that Q has the same eigenfunctions as the linear operator A and we have diagonalized the operator A by a Fourier spectral method then (2.11) reduces to an Ornstein–Uhlenbeck process in each Fourier mode as in [12]. This can then be simulated numerically with the correct mean and variance in each mode. The new scheme (2.10) uses a finite element discretization and projects the linear functional of the noise and hence we expect superior approximation properties over a standard semi-implicit Euler discretization for a finite element discretization.

The existence and uniqueness of the solution of equation is well known ([13, 5, 18, 4]). For convergence proofs below we need sufficient regularity of the mild solution X, and therefore we will use some weak assumptions. We describe in detail the weak assumptions that we make on the linear operator A, the nonlinear term F and the noise dW.

Assumption 2.2. ***[Linear operator]*** *The linear operator* $-A$ *given in* (2.1) *is positive definite. Then there exists sequences of positive real eigenvalues* $\{\lambda_n\}_{n\in\mathbb{N}^d}$

with $\inf_{i\in\mathbb{N}^d}\lambda_i > 0$ *and an orthonormal basis in* H *of eigenfunctions* $\{e_i\}_{i\in\mathbb{N}^d}$ *such that the linear operator* $-A : \mathcal{D}(-A) \subset H \to H$ *is represented as*

$$-Av = \sum_{i\in\mathbb{N}^d} \lambda_i (e_i, v) e_i \qquad \forall \quad v \in \mathcal{D}(-A)$$

where the domain of $-A$, $\mathcal{D}(-A) = \{v \in H : \sum_{i\in\mathbb{N}^d} \lambda_i^2 |(e_i, v)| < \infty\}$.

Assumption 2.3. ***[Nonlinearity]***
We assume that there exists a positive constant $L > 0$ *such that* F *satisfies one of the following.*
(a) Let $\mathcal{V}$ *be a separable Banach space such that* $\mathcal{D}((-A)^{1/2}) \subset \mathcal{V} \subset H = L^2(\Omega)$ *continuously.*
$F : \mathcal{V} \to \mathcal{V}$ *is a twice continuously Nemytskii Fréchet differentiable mapping with*

$$\|F'(v)w\| \leq L\|w\|, \qquad \|F'(v)\|_{L(\mathcal{V})} \leq L, \qquad \|F''(v)\|_{L^{(2)}(\mathcal{V})} \leq L$$

and

$$\|(F'(u))^*\|_{L(\mathcal{D}((-A)^{1/2}))} \leq L(1 + \|u\|_{\mathcal{D}((-A)^{1/2})}) \qquad \forall \quad v, w \in \mathcal{V}, \quad u \in \mathcal{D}((-A)^{1/2}),$$

where $(F'(u))^*$ *is the adjoint of* $F'(u)$ *defined by*

$$((F'(u))^* v, w) = (v, F'(u)w) \qquad \forall \quad v, w \in H = L^2(\Omega).$$

As a consequence

$$\|F(Z) - F(Y)\| \leq L\|Z - Y\| \qquad \forall\, Z, Y \in H,$$

and $\forall\, Y \in H = L^2(\Omega)$

$$\|F(Y)\| \leq \|F(0)\| + \|F(Y) - F(0)\| \leq \|F(0)\| + L\|Z\| \leq C(\|F(0)\| + \|Y\|).$$

(b) F *is globally Lipschitz continuous from* $(H^1(\Omega), \|.\|_{H^1(\Omega)})$ *to* $(H = L^2(\Omega), \|.\|)$, *i.e,*

$$\|F(Z) - F(Y)\| \leq L\|Z - Y\|_{H^1(\Omega)} \qquad \forall Z, Y \in H^1(\Omega).$$

Remark 2.4. *As the* $\mathbb{R}-$ *Hilbert space is* $H = L^2(\Omega)$, $F'(u)$ *is self-adjoint we have* $F'(u)^* = F'(u)$, $u \in \mathcal{D}((-A)^{1/2})$. *This come obviously to the fact that*

$$(v, w) = \int_\Omega vw\, dx \qquad v, w \in L^2(\Omega).$$

We keep the adjoint notation $F'(u)^*$ *in Assumption 2.3 as results in this paper can be generalized in any Hilbert space* H. *Possibly expression of* F *is*

$$F(u)(x) = f(x, u(x)) \quad x \in \Omega,$$

where $f : \Omega \times \mathbb{R} \to \mathbb{R}$ *is twice continuously differentiable function with the bounded partial derivatives.*

We note that in condition (a) of Assumption 2.3, the possible choice of $\mathcal{V}$ can be H or $H^1(\Omega)$. Strong convergence proofs with a Nemytskii operator for the nonlinearity were recently obtained in [13].

We now turn our attention to the noise term and introduce spaces and notation that we need to define the Q-Wiener process. An operator $T \in L(H)$ is Hilbert-Schmidt if

$$\|T\|_{HS}^2 := \sum_{i\in\mathbb{N}^d} \|Te_i\|^2 < \infty.$$

The sum in $\|.\|_{HS}^2$ is independent of the choice of the orthonormal basis in H. We denote the space of Hilbert- Schmidt operator from $Q^{1/2}(H)$ to H by $L_2^0 := HS(Q^{1/2}(H), H)$ and the corresponding norm $\|.\|_{L_2^0}$ by

$$\|T\|_{L_2^0} := \|TQ^{1/2}\|_{HS} = \left(\sum_{i\in\mathbb{N}^d} \|TQ^{1/2}e_i\|^2\right)^{1/2} \quad T \in L_2^0.$$

Let φ be a continuous L_2^0-process. We have the following equality known as the Ito's isometry

$$\mathbf{E}\|\int_0^t \varphi dW\|^2 = \int_0^t \mathbf{E}\|\varphi\|_{L_2^0}^2 ds = \int_0^t \mathbf{E}\|\varphi Q^{1/2}\|_{HS}^2 ds.$$

We assume sufficient regularity of the noise for the existence of a mild solution and to project the noise into the space V_h. To be specific we need that the stochastic process O is in H^1 or H^2 in space. This is used in our proofs to apply the finite element projection P_h in the errors estimates.

Notice that for all $t \in [0, T]$ the process $O(t)$ is an adapted stochastic process to the filtration $(\mathcal{F}_t)_{t\geq 0}$ with continuous sample paths such that $O(t_2) - S(t_2 - t_1)O(t_1),\ 0 \leq t_1 < t_2 \leq T$ is independent of $\mathcal{F}_{t_1}$.

Assumption 2.5. ***[Noise and stochastic process** O]* *We assume one or both of the following.*

(a) The sequence (q_i) *in the noise representation* (1.2) *satisfies*

$$\sum_{i\in\mathbb{N}^d} \lambda_i^{r-1} q_i < \infty \Leftrightarrow \|(-A)^{(r-1)/2} Q^{1/2}\|_{HS} < \infty \quad r \in \{1, 2\}.$$

As a consequence

$$O(t) \in L_2(\mathbb{D}, \mathcal{D}((-A)^{r/2})), \qquad 0 \leq t \leq T \tag{2.12}$$

$$\mathcal{D}((-A)^{r/2}) = \mathbb{H} \cap H^r(\Omega),\ r \in \{1, 2\}. \tag{2.13}$$

(b) For some $\theta \in (0, 1/2]$ *and positive constant* $C > 0$

$$\mathbf{E}\left(\|O(t_2) - O(t_1)\|_{\mathcal{V}}^4\right) \leq C(t_2 - t_1)^{4\theta} \qquad 0 \leq t_1 < t_2 \leq T,$$

where $\mathcal{V}$ *is defined in Assumption 2.3(a).*

We note again that we have assumed only for convenience that Q and A have the same eigenfunctions. In particular it is useful to characterize the decay of q_i in the condition (a) of Assumption 2.3 above.

One can prove that in condition (b) of Assumption 2.5, for $\mathcal{V} = H = L^2(\Omega)$ we can take $\theta = 1/2$. For $\mathcal{V} = H^1(\Omega)$ we can take $\theta = 1/2$ if $O(t) \in L_2(\mathbb{D}, \mathcal{D}(-A)),\ 0 \leq t \leq T$

and $\theta \neq 1/2$ but close to $1/2$ if $O(t) \in L_2(\mathbb{D}, \mathcal{D}((-A)^{1/2}))$, $0 \leq t \leq T$. The proof is similar for that in Lemma 3.2.

Remark 2.6. *It is important to notice that for* $0 \leq \gamma \leq 1$, *if* $X_0 \in L_2(\mathbb{D}, \mathcal{D}((-A)^{\gamma}))$ *with* $\mathbf{E}\|(-A)^{\gamma} X_0)\|^4 < \infty$, *Assumption 2.2, Assumption 2.3 and Assumption 2.5 ensure the existence of the unique solution* $X(t) \in L_2(\mathbb{D}, \mathcal{D}((-A)^{\gamma}))$ *such that*

$$\mathbf{E}\left(\sup_{0\leq s\leq T} \|(-A)^{\gamma} X(s))\|^4\right) < \infty.$$

In general if $X_0 \in L_2(\mathbb{D}, \mathcal{D}((-A)^{\gamma}))$ *with* $\mathbf{E}\|(-A)^{\gamma} X_0)\|^4 < \infty$, *and* $O(t) \in L_2(\mathbb{D}, \mathcal{D}((-A)^{\alpha}))$ *then* $X \in L_2(\mathbb{D}, \mathcal{D}((-A)^{min(\gamma,\alpha)}))$ *with*

$$\mathbf{E}\left(\sup_{0\leq s\leq T} \|(-A)^{min(\gamma,\alpha)} X(s)\|^4\right) < \infty.$$

More information about properties of the solution of the SPDE (1.1) can be found in [13].

Assumption 2.7. ***[Initial solution*** X_0 ***and extra assumption for nonlinearity*** *F]*

- *We assume that when initial data* $X_0 \in L_2(\mathbb{D}, \mathcal{D}((-A)^{\gamma}))$ *then* $\mathbf{E}\|(-A)^{\gamma} X_0\|^4 < \infty$, *with* $0 \leq \gamma \leq 1$.
- *Further we assume in our main results that for a solution* $X(t)$ *of our SPDE* (1.1) $F(X(t)) \in L_2(\mathbb{D}, \mathbb{H}), \forall\, t \in [0,T]$ *with* $\sup_{0\leq s\leq T} \mathbf{E}\|F(X(s))\|_1^2 < \infty$.

Remark 2.8. *It is important to notice that we denote by Assumption 2.3(i) and Assumption 2.5(i) to specify condition (i) of Assumption 2.3 and Assumption 2.5,* $i \in \{a, b\}$.

2.1. Main results. Throughout the article we let N be the number of terms of truncated noise, $\mathcal{I}_N = \{1, 2, ..., N\}^d$ and take $t_m = m\Delta t \in (0, T]$, where $T = M\Delta t$ for $m, M \in \mathbb{N}$. We take C to be a constant that may depend on T and other parameters but not on Δt, N or h.

Our first result is a strong convergence result in L^2 when the non-linearity satisfies the Lipschitz condition of Assumption 2.3 (a).

Theorem 2.9. *Suppose that Assumptions 2.2, 2.3(a),2.5 (a) and (b) (with* $r = 1, 2$*) and Assumption 2.7 are satisfied. Let* $X(t_m)$ *be the mild solution of equation* (1.1) *represented by (1.3) and* X_m^h *be the numerical approximation through* (2.10). *Let* $1/2 \leq \gamma < 1$ *and set* $\sigma = \min(2\theta, \gamma)$, *and let* $\theta \in (0, 1/2]$ *be defined as in Assumption 2.5. If* $X_0 \in L_2(\mathbb{D}, \mathcal{D}((-A)^{\gamma}))$ *then*

$$\left(\mathbf{E}\|X(t_m) - X_m^h\|^2\right)^{1/2} \leq C\left(t_m^{-1/2}(h^r + \Delta t^{\sigma}) + \Delta t\,|\ln(\Delta t)| + \left(\inf_{j\in\mathbb{N}^d\setminus\mathcal{I}_N} \lambda_j\right)^{-r/2}\right).$$

If $X_0 \in L_2(\mathbb{D}, \mathcal{D}(-A))$ *then*

$$\left(\mathbf{E}\|X(t_m) - X_m^h\|^2\right)^{1/2} \leq C\left(h^r + \Delta t^{2\theta} + \Delta t\,|\ln(\Delta t)| + \left(\inf_{j\in\mathbb{N}^d\setminus\mathcal{I}_N} \lambda_j\right)^{-r/2}\right).$$

If we assume stronger regularity on the noise we can obtain a strong error estimate in the H^1 norm.

THEOREM 2.10. *Suppose that Assumptions 2.2, 2.3(b), 2.5(a) (with $r = 2$) and Assumption 2.7 are satisfied. Let X be the solution mild of equation (1.1) represented by equation (1.3). Then we have the following:*
If $X_0 \in L_2(\mathbb{D}, \mathcal{D}(-A))$ then

$$(\mathbf{E}\|X(t_m) - X^h_m\|^2_{H^1(\Omega)})^{1/2} \leq C \left((h + \Delta t^{1/2-\epsilon} t_m^{-1/2}) + \left(\inf_{j\in\mathbb{N}^d\setminus\mathcal{I}_N} \lambda_j \right)^{-1/2} \right).$$

If $X_0 \in L_2(\mathbb{D}, \mathcal{D}(-A))$ and $-AX_0 \in L_2(\mathbb{D}, H^1(\Omega))$ then

$$(\mathbf{E}\|X(t_m) - X^h_m\|^2_{H^1(\Omega)})^{1/2} \leq C \left((h + \Delta t^{1/2-\epsilon}) + \left(\inf_{j\in\mathbb{N}^d\setminus\mathcal{I}_N} \lambda_j \right)^{-1/2} \right),$$

for small enough $\epsilon \in (0, 1/2)$.
First note in both theorems we see that if the initial data are not sufficiently smooth then the error is dominated by this term. This behaviour is typical of non-smooth error estimates. Secondly we remark that if we denote by N_h the number of vertices in the finite element mesh then it is well known (see for example [22]) that if $N \geq N_h$ then

$$\left(\inf_{j\in\mathbb{N}^d\setminus\mathcal{I}_N} \lambda_j \right)^{-1} \leq Ch^2 \qquad \text{and} \qquad \left(\inf_{j\in\mathbb{N}^d\setminus\mathcal{I}_N} \lambda_j \right)^{-1/2} \leq Ch.$$

As a consequence the estimates in Theorem 2.9 and Theorem 2.10 can be expressed as a function of h and Δt only and it is the error from the finite element approximation that dominates. If $N < N_h$ then it is the error from the projection P_N of the noise onto a finite number of modes that dominates
From Theorem 2.10 we also get an estimate in the root mean square $L^2(\Omega)$ norm in the case that the nonlinear function F satisfies Assumption 2.3 (b).
Finally we note that if the nodes of the finite element mesh coincide with evaluations of $O(x, t)$ then the projection operator P_h is trivial. This also leads to a computational advantage as we no longer need the projection, we comment further in Section 4.

3. Proofs of main results.

3.1. Some preparatory results. We introduce the Riesz representation operator $R_h : V \to V_h$ defined by

$$(-AR_hv, \chi) = (-Av, \chi) = a(v, \chi) \qquad v \in V,\ \forall \chi \in V_h. \tag{3.1}$$

Under the regularity assumptions on the triangulation and in view of the $V-$ellipticity (2.2), it is well known (see [7]) that the following error bounds holds

$$\|R_hv - v\| + h\|R_hv - v\|_{H^1(\Omega)} \leq Ch^r\|v\|_{H^r(\Omega)}, \quad v \in V \cap H^r(\Omega),\ r \in \{1, 2\} \tag{3.2}$$

It follows that

$$\|P_hv - v\| \leq Ch^r\|v\|_{H^r(\Omega)} \quad \forall v \in V \cap H^r(\Omega), \quad r \in \{1, 2\}. \tag{3.3}$$

We examine the deterministic linear problem. Find $u \in V$ such that

$$u' = Au \qquad \text{given} \quad u(0) = v, \quad t \in (0, T]. \tag{3.4}$$

The corresponding semi–discretization in space is : find $u_h \in V_h$ such that $u_h' = A_h u_h$ where $u_h^0 = P_h v$. The full discretization of (3.4) using implicit Euler in time is given by

$$u_h^{n+1} = (\mathrm{I} - \Delta t\ A_h)^{-(n+1)}\ P_h v = S_{h,\Delta t}^{n+1}\ P_h v.$$

We consider the error at $t_n = n\Delta t$ and define the operator T_n from

$$u(t_n) - u_h^n = (S(t_n) - (I - \Delta t\, A_h)^{-n}\ P_h)v := T_n v. \tag{3.5}$$

LEMMA 3.1. *The following estimates hold on the numerical approximation to (3.4).* ***Estimation in*** $H = L^2(\Omega)$ ***norm****. If $v \in \mathbb{H}$ then*

$$\|u(t_n) - u_h^n\| = \|T_n v\| \leq C t_n^{-1/2}(h^2 + \Delta t)\|v\|_1 \tag{3.6}$$

and if $v \in \mathcal{D}(-A) = \mathbb{H} \cap H^2(\Omega)$

$$\|u(t_n) - u_h^n\| = \|T_n v\| \leq C(h^2 + \Delta t)\|v\|_2. \tag{3.7}$$

Estimation in $H^1(\Omega)$ ***norm****. If $v \in \mathbb{H}$ then*

$$\|u(t_n) - u_h^n\|_{H^1(\Omega)} = \|T_n v\|_{H^1(\Omega)} \leq C\|v\|_1(t_n^{-1/2}h + t_n^{-1}\Delta t). \tag{3.8}$$

If $v \in \mathcal{D}(-A) = \mathbb{H} \cap H^2(\Omega)$

$$\|u(t_n) - u_h^n\|_{H^1(\Omega)} = \|T_n v\|_{H^1(\Omega)} \leq C\|v\|_2(h + t_n^{-1/2}\Delta t). \tag{3.9}$$

Finally, if $v \in \mathcal{D}(-A)$ and $Av \in H^1(\Omega)$ then

$$\|u(t_n) - u_h^n\|_{H^1(\Omega)} = \|T_n v\|_{H^1(\Omega)} \leq C\| - Av\|_{H^1(\Omega)}(h + \Delta t). \tag{3.10}$$

Proof. We give here some references for the proof. Estimates in the $H = L^2(\Omega)$ norm are given in [7, 21]. In [21], $A = \Delta$ with Dirichlet boundary conditions. Estimates in the $H^1(\Omega)$ norm are the special cases of Theorem 5.3 in [16] where the proof is given for a general semi–linear parabolic problem with a locally Lipschitz nonlinear term. To obtain our result from [16] note that $u(t) = S(t)v$ so that we have the analogue of [16, Theorem 5.2]

$$\|u_t(t)\|_{H^2(\Omega)} + \|u_{tt}(t)\| \leq Ct^{-3/2}\|v\|_1 \quad \text{if} \quad v \in \mathbb{H},$$
$$\|u_t(t)\|_{H^2(\Omega)} + \|u_{tt}(t)\| \leq Ct^{-1}\|v\|_2 \quad \text{if} \quad v \in \mathcal{D}(-A) = V \cap H^2(\Omega),$$
$$\|u_t(t)\|_{H^2(\Omega)} + \|u_{tt}(t)\| \leq Ct^{-1/2}\| - Av\|_{H^1(\Omega)} \quad \text{if} \quad v \in \mathcal{D}(-A)\,\text{and}\, - Av \in H^1(\Omega).$$

Using these in the proof of [16, Theorem 5.3] gives the result. □
Our second preliminary lemma concerns the mild solution SPDE of (1.1).
LEMMA 3.2. *Let X be the mild solution of (1.1) given in (1.3), let $0 < \gamma < 1$ and $t_1, t_2 \in [0,T], \quad t_1 < t_2$.*
(i) If $X_0 \in L_2(\mathbb{D}, \mathcal{D}((-A)^\gamma))$, $\|(-A)^{(\alpha-1)/2}Q^{1/2}\|_{HS} < \infty$ with $0 < \alpha \leq 2$ and suppose F satisfies Assumption 2.3 (a). Set $\sigma = \min(\gamma, 1/2, (\alpha - \epsilon)/2)$ with $\epsilon \in (0, \alpha/2)$ very small enough, then

$$\mathbf{E}\|X(t_2) - X(t_1)\|^2 \leq C(t_2 - t_1)^{2\sigma}\left(\mathbf{E}\|X_0\|_\gamma^2 + \mathbf{E}\left(\sup_{0\leq s\leq T}(\|F(0)\| + \|X(s)\|)\right)^2 + 1\right).$$

Furthermore

$$\mathbf{E}\|(X(t_2)-O(t_2))-(X(t_1)-O(t_1))\|^2 \leq C(t_2-t_1)^{2\gamma} \quad 0\leq\gamma\leq 1.$$

(ii) If $X_0 \in L_2(\mathbb{D},\mathcal{D}((-A)^{(\gamma+1)/2}))$, $\|(-A)^{1/2}Q^{1/2}\|_{HS}<\infty$ *and* $F(X(t)) \in L_2(\mathbb{D},H^1(\Omega))$, $\forall\ t\in[0,T]$ *with* $\mathbf{E}\left(\sup\limits_{0\leq s\leq T}\|F(X(s))\|_{H^1(\Omega)}\right)^2<\infty$ *then*

$$\mathbf{E}\|X(t_2)-X(t_1)\|^2 \leq C(t_2-t_1)^{\gamma}\left(\mathbf{E}\|X_0\|^2_{(\gamma+1)}+\mathbf{E}\left(\sup_{0\leq s\leq T}\|F(X(s))\|_{H^1(\Omega)}\right)^2+1\right).$$

Proof.
Proof of the first claim of part (i) The proof is the same as in [20, Lemma 5.10] with the weak assumption $\|(-A)^{\alpha/2}Q^{1/2}\|_{HS}<\infty$.
Consider the difference

$$\begin{aligned}
&X(t_2)-X(t_1)\\
&=(S(t_2)-S(t_1))\,X_0+\left(\int_0^{t_2}S(t_2-s)F(X(s))ds-\int_0^{t_1}S(t_1-s)F(X(s))ds\right)\\
&\quad+\left(\int_0^{t_2}S(t_2-s)dW(s)-\int_0^{t_1}S(t_1-s)dW(s)\right)\\
&=I+II+III
\end{aligned}$$

so that $\mathbf{E}\|X(t_2)-X(t_1)\|^2\leq 3(\mathbf{E}\|I\|^2+\mathbf{E}\|II\|^2+\mathbf{E}\|III\|^2)$.
We estimate each of the terms I, II and III. For I, using Proposition 2.1 yields

$$\|I\|=\|S(t_1)(-A)^{-\gamma}(\mathrm{I}-S(t_2-t_1))(-A)^{\gamma}X_0\| \quad\leq\quad C(t_2-t_1)^{\gamma}\|X_0\|_{\gamma}.$$

Then $\mathbf{E}\|I\|^2\leq C(t_2-t_1)^{2\gamma}\mathbf{E}\|X_0\|^2_{\gamma}$. For the term II, we have

$$\begin{aligned}
II&=\int_0^{t_1}(S(t_2-s)-S(t_1-s))F(X(s))ds+\int_{t_1}^{t_2}S(t_2-s)F(X(s))ds\\
&=II_1+II_2.
\end{aligned}$$

We now estimate each term II_1 and II_2. For II_1

$$\begin{aligned}
\mathbf{E}\|II_1\|^2&=\mathbf{E}\|\int_0^{t_1}(S(t_2-s)-S(t_1-s))F(X(s))ds\|^2\\
&\leq\mathbf{E}\left(\int_0^{t_1}\|(S(t_2-s)-S(t_1-s))F(X(s))\|ds\right)^2\\
&\leq\left(\int_0^{t_1}\|(S(t_2-s)-S(t_1-s))\|_{L(L^2(\Omega))}ds\right)^2\mathbf{E}\left(\sup_{0\leq s\leq T}\|F(X(s))\|\right)^2.
\end{aligned}$$

For $0 \leq \gamma < 1$, Proposition 2.1 yields

$$
\begin{aligned}
\mathbf{E}\|II_1\|^2 &\leq \left(\int_0^{t_1} \|S(t_1-s)(-A)^{\gamma}(-A)^{-\gamma}(\mathrm{I}-S(t_2-t_1))\|_{L(L^2(\Omega))}ds\right)^2 \mathbf{E}\left(\sup_{0\leq s\leq T}\|F(X(s))\|\right)^2\\
&\leq \left(\int_0^{t_1} \|(-A)^{\gamma}S(t_1-s)(-A)^{-\gamma}(\mathrm{I}-S(t_2-t_1))\|_{L(L^2(\Omega))}ds\right)^2 \mathbf{E}\left(\sup_{0\leq s\leq T}\|F(X(s))\|\right)^2\\
&\leq C(t_2-t_1)^{2\gamma}\left(\int_0^{t_1}(t_1-s)^{-\gamma}ds\right)^2 \mathbf{E}\left(\sup_{0\leq s\leq T}\|F(X(s))\|\right)^2\\
&\leq C(t_2-t_1)^{2\gamma}\,\mathbf{E}\left(\sup_{0\leq s\leq T}\|F(X(s))\|\right)^2.
\end{aligned}
$$

For II_2, using the fact that the semigroup is bounded, we have

$$
\begin{aligned}
\mathbf{E}\|II_2\|^2 &= \mathbf{E}\|\int_{t_1}^{t_2} S(t_2-s)F(X(s))ds\|^2\\
&\leq \mathbf{E}\left(\int_{t_1}^{t_2}\|S(t_2-s)F(X(s))\|ds\right)^2\\
&\leq \mathbf{E}\left(\int_{t_1}^{t_2}\|F(X(s))\|ds\right)^2 \quad\leq\quad C(t_2-t_1)^2\mathbf{E}\left(\sup_{0\leq s\leq T}\|F(X(s))\|\right)^2.
\end{aligned}
$$

Hence, if F satisfies Assumption 2.3 (a) we have

$$
\mathbf{E}\|II\|^2 \leq 2(\mathbf{E}\|II_1\|^2+\mathbf{E}\|II_2\|^2) \leq C(t_2-t_1)^{2\gamma}\mathbf{E}\left(\sup_{0\leq s\leq T}(\|F(0)\|+\|X(s)\|)\right)^2.
$$

For term III we have

$$
III = \int_0^{t_1}\left(S(t_2-s)-S(t_1-s)\right)dW(s) + \int_{t_1}^{t_2} S(t_2-s)dW(s) \quad=\quad III_1+III_2.
$$

First using the Ito isometry property and then $0<\alpha\leq 2$ we have

$$
\begin{aligned}
\mathbf{E}\|III_1\|^2 &= \mathbf{E}\|\int_0^{t_1}\left(S(t_2-s)-S(t_1-s)\right)dW(s)\|^2\\
&= \int_0^{t_1}\mathbf{E}\|\left(S(t_2-s)-S(t_1-s)\right)Q^{1/2}\|_{HS}^2ds\\
&= \int_0^{t_1}\mathbf{E}\|(S(t_2-s)-S(t_1-s))(-A)^{-(\alpha-1)/2}(-A)^{(\alpha-1)/2}Q^{1/2}\|_{HS}^2ds.
\end{aligned}
$$

Using Proposition 2.1, that $\|(-A)^{(\alpha-1)/2}Q^{1/2}\|_{HS}<\infty$ and boundedness of S yields

$$
\begin{aligned}
\mathbf{E}\|III_1\|^2 &\leq \int_0^{t_1}\|(-A)^{-(\alpha-1)/2}(S(t_2-s)-S(t_1-s))\|_{L(L^2(\Omega))}^2ds\\
&= \int_0^{t_1}\|(-A)^{(1-\epsilon)/2}S(t_1-s)(-A)^{-(\alpha-\epsilon)/2}(\mathbf{I}-S(t_2-t_1))\|_{L(L^2(\Omega))}^2ds\\
&\leq C(t_2-t_1)^{\alpha-\epsilon}\int_0^{t_1}(t_1-s)^{1-\epsilon}ds\\
&\leq C(t_2-t_1)^{\alpha-\epsilon},
\end{aligned}
$$

with $\epsilon \in (0, \alpha/2)$ small enough.
Let us estimate $\mathbf{E}\|III_2\|$. Using the Ito isometry again, and that for $0 < \alpha \leq 2$ we assume $\|(-A)^{(\alpha-1)/2}Q^{1/2}\|_{HS} < \infty$ then $\|Q^{1/2}\|_{HS} < \infty$ (by taking $\alpha = 1$), with boundedness of S yields

$$\mathbf{E}\|III_2\|^2 = \mathbf{E}\|\int_{t_1}^{t_2} S(t_2-s)dW(s)\|^2 \quad = \quad \int_{t_1}^{t_2} \|S(t_2-s)Q^{1/2}\|^2_{HS}ds \leq \ C(t_2-t_1).$$

Hence

$$\mathbf{E}\|III\|^2 \leq 2(\mathbf{E}\|III_1\|^2 + \mathbf{E}\|III_2\|^2) \leq C(t_2-t_1)^{2\min(\gamma,(\alpha-\epsilon)/2,1/2)} = C(t_2-t_1)^{2\sigma}.$$

with $\epsilon \in (0, \alpha/2)$ small enough. Combining our estimates of $\mathbf{E}\|I\|^2, \mathbf{E}\|II\|^2$ and $\mathbf{E}\|III\|^2$ ends the first part of the first claim in the lemma.
The proof of the second claim of part (i) and part (ii) can be found in [20, Lemma 5.10]. □

3.2. Proof of Theorem 2.9. We now estimate $\left(\mathbf{E}\|X(t_m) - X^h_m\|^2\right)^{1/2}$. Again we look at the difference between the mild solution and our numerical approximation (2.10). By construction of the approximation from (2.9) and (2.8) we have that

$$\begin{aligned} X(t_m) - X^h_m &= \overline{X}(t_m) + O(t_m) - X^h_m \\ &= \overline{X}(t_m) + O(t_m) - \left(Z^h_m + P_hP_NO(t_m)\right) \\ &= \left(\overline{X}(t_m) - Z^h_m\right) + (P_N(O(t_m)) - P_hP_N(O(t_m))) + (O(t_m) - P_N(O(t_m))) \\ &= I + II + III, \end{aligned} \tag{3.11}$$

where $\overline{X}(t)$ is given by

$$\overline{X}(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))ds, \qquad t \in [0,T].$$

and Z^h_m by (2.8).
Then $\left(\mathbf{E}\|X(t_m) - X^h_m\|^2\right)^{1/2} \leq \left(\mathbf{E}\|I\|^2\right)^{1/2} + \left(\mathbf{E}\|II\|^2\right)^{1/2} + \left(\mathbf{E}\|III\|^2\right)^{1/2}$ and we estimate each term. Since the first term will require the most work we first estimate the other two.
Let us estimate $\left(\mathbf{E}\|II\|^2\right)^{1/2}$. To do this we use the finite element estimate (3.3), the regularity of the noise from Assumption 2.5 (a) and the fact that P_N is bounded. Then for $r \in \{1,2\}$ we have

$$\mathbf{E}\|II\|^2 \leq Ch^{2r}\mathbf{E}\|P_N(O(t_m))\|_{H^r(\Omega)} \leq Ch^{2r}\mathbf{E}\|O(t_m)\|_{H^r(\Omega)}.$$

Using $\|.\|_{H^r(\Omega)} \equiv \|(-A)^{r/2}.\|$ in $\mathcal{D}((-A)^{r/2})$ and the Ito isometry yields

$$\begin{aligned} \mathbf{E}\|II\|^2 &\leq Ch^{2r}\mathbf{E}\|(-A)^{r/2}\int_0^{t_m} S(t_m-s)dW(s)\|^2 \\ &= Ch^{2r}\int_0^{t_m} \|(-A)^{r/2}S(t_m-s)Q^{1/2}\|^2_{HS}ds \\ &= Ch^{2r}\int_0^{t_m} \sum_{i\in\mathbb{N}^d} \lambda_i^r q_i e^{-2\lambda_i(t_m-s)}ds \\ &= \frac{Ch^{2r}}{2}\sum_{i\in\mathbb{N}^d}\lambda_i^{r-1}q_i(1-e^{-\lambda_i t_m}) \leq \frac{Ch^{2r}}{2}\sum_{i\in\mathbb{N}^d}\lambda_i^{r-1}q_i = \frac{Ch^{2r}}{2}\|(-A)^{(r-1)/2}Q^{1/2}\|_{HS}. \end{aligned}$$

Thus, since the noise is in H^r, using relation (2.12) of Assumption 2.5 yields $(\mathbf{E}\|II\|^2)^{1/2} \leq Ch^r$.
For the third term III

$$\mathbf{E}\|III\|^2 = \mathbf{E}\|(\mathrm{I}-P_N)O(t_m)\|^2 = \mathbf{E}\|(\mathrm{I}-P_N)(-A)^{-r/2}(-A)^{r/2}O(t_m)\|^2,$$

and so

$$\mathbf{E}\|III\|^2 \leq \|(\mathrm{I}-P_N)(-A)^{-r/2}\|^2_{L(L^2(\Omega))}\mathbf{E}\|(-A)^{r/2}O(t_m)\|^2 \leq C\left(\inf_{j\in\mathbb{N}^d\setminus\mathcal{I}_N}\lambda_j\right)^{-r}.$$

We now turn our attention to the first term $\mathbf{E}\|I\|^2$. Using the definition of $S_{h,\Delta t}$ in (2.7) we can write (2.8) as

$$Z^h_m = S^m_{h,\Delta t}P_hX_0 + \Delta t\sum_{k=0}^{m-1}S^{(m-k)}_{h,\Delta t}P_hF(Z^h_k + P_hP_NO(t_k)).$$

Then using the definition of T_m from (3.5) the first term I can be expanded

$$\begin{aligned}
I &= T_mX_0 + \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} S(t_m-s)F(X(s)) - S^{(m-k)}_{h,\Delta t}P_hF(Z^h_k + P_hP_NO(t_k))ds\\
&= T_mX_0 + \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} S^{(m-k)}_{h,\Delta t}P_h(F(X(t_k)) - F(Z^h_k + P_hP_NO(t_k))))ds\\
&\quad+\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} S^{(m-k)}_{h,\Delta t}P_h(F(X(s)) - F(X(t_k)))ds\\
&\quad+\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} (S(t_m-t_k) - S^{(m-k)}_{h,\Delta t}P_h)F(X(s))ds\\
&\quad+\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} (S(t_m-s) - S(t_m-t_k))F(X(s))ds\\
&= I_1 + I_2 + I_3 + I_4 + I_5. \qquad (3.12)
\end{aligned}$$

Then

$$\left(\mathbf{E}\|I\|^2\right)^{1/2} \leq \left(\mathbf{E}\|I_1\|^2\right)^{1/2} + \left(\mathbf{E}\|I_2\|^2\right)^{1/2} + \left(\mathbf{E}\|I_3\|^2\right)^{1/2} + \left(\mathbf{E}\|I_4\|^2\right)^{1/2} + \left(\mathbf{E}\|I_5\|^2\right)^{1/2}.$$

For I_1, if $X_0 \in L_2(\mathbb{D}, \mathcal{D}((-A)^\gamma))$, $1/2 \leq \gamma < 1$, equation (3.6) of Lemma 3.1 gives

$$(\mathbf{E}\|I_1\|^2)^{1/2} \leq C(t_m^{-1/2}(h^2+\Delta t))\left(\mathbf{E}\|X_0\|^2_1\right)^{1/2}$$

and if $X_0 \in L_2(\mathbb{D}, \mathcal{D}(-A))$, equation (3.7) of Lemma 3.1 gives

$$(\mathbf{E}\|I_1\|^2)^{1/2} \leq C(h^2+\Delta t)\left(\mathbf{E}\|X_0\|^2_2\right)^{1/2}.$$

If F satisfies Assumption 2.3 (a), then using the Lipschitz condition, triangle inequality and the fact that $S^{(m-k)}_{h,\Delta t}$ and P_h are an bounded operators, we have

$$\begin{aligned}
\left(\mathbf{E}\|I_2\|^2\right)^{1/2} &\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\left(\mathbf{E}\|F(X(t_k)) - F((Z^h_k + P_hP_NO(t_k))\|^2\right)^{1/2}ds\\
&\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\left(\mathbf{E}\|X(t_k) - X^h_k\|^2\right)^{1/2}ds.
\end{aligned}$$

Let us estimate $\left(\mathbf{E}\|I_3\|^2\right)^{1/2}$. We add in and subtract out $O(s)$ and $O(t_k)$

$$\begin{aligned}
\left(\mathbf{E}\|I_3\|^2\right)^{1/2} &= \left(\mathbf{E}\|\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{(m-k)} P_h \left(F(X(s)) - F(X(t_k))\right) ds\|^2\right)^{1/2} \\
&\leq \left(\mathbf{E}\|\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{(m-k)} P_h \left(F(X(s)) - F(X(t_k)+O(s)-O(t_k)))\right) ds\|^2\right)^{1/2} \\
&\quad + \left(\mathbf{E}\|\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{(m-k)} P_h \left(F(X(t_k)+O(s)-O(t_k)) - F(X(t_k)))\right) ds\|^2\right)^{1/2} \\
&:= \left(\mathbf{E}\|I_3^1\|^2\right)^{1/2} + \mathbf{E}\left(\|I_3^2\|^2\right)^{1/2}.
\end{aligned}$$

Applying the Lipschitz condition in Assumption 2.5, using the fact the semigroup is bounded and according to Lemma 3.2, for $X_0 \in L_2(\mathbb{D}, \mathcal{D}((-A)^{\gamma}))$, $\quad 0 \leq \gamma \leq 1$ we therefore have

$$\begin{aligned}
\left(\mathbf{E}\|I_3^1\|^2\right)^{1/2} &\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} \left(\mathbf{E}\|(X(s)-O(s)) - (X(t_k)-O(t_k))\|^2\right)^{1/2} ds \\
&\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} (s-t_k)^{\gamma} ds \quad \leq \quad C\Delta t^{\gamma}.
\end{aligned}$$

Let us now estimate $\mathbf{E}\left(\|I_3^2\|^2\right)^{1/2}$. The analysis below follows the same steps as in [13] although the approximating semigroup $S_{h,\Delta t}$ is different here. Applying a Taylor expansion to F gives

$$\mathbf{E}\left(\|I_3^2\|^2\right)^{1/2} \leq I_3^{21} + I_3^{22} + I_3^{23},$$

with

$$\begin{aligned}
I_3^{21} &= \left(\mathbf{E}\|\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{(m-k)} P_h F'(X(t_k))(O(s) - S(s-t_k)O(t_k)) ds\|^2\right)^{1/2} \\
I_3^{22} &= \left(\mathbf{E}\|\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{(m-k)} P_h F'(X(t_k))(S(s-t_k)O(t_k) - O(t_k)) ds\|^2\right)^{1/2} \\
I_3^{23} &= \left(\mathbf{E}\|\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{(m-k)} \int_0^1 G(1-r) dr ds\|^2\right)^{1/2},
\end{aligned}$$

where G is the expression $G := P_h F''(X(t_k)) + r(O(s)-O(t_k))(O(s)-O(t_k), O(s) - O(t_k))$. Using the fact that $O(t_2) - S(t_2-t_1)O(t_1)$, $0 \leq t_1 < t_2 \leq T$ is independent of $\mathcal{F}_{t_1}$, one can show, as in [13], that

$$\left(I_3^{21}\right)^2 = \sum_{k=0}^{m-1}\mathbf{E}\|\int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{(m-k)} P_h F'(X(t_k))(O(s) - S(s-t_k)O(t_k)) ds\|^2.$$

Therefore as $S_{h,\Delta t}$ is bounded we have

$$
\begin{aligned}
I_3^{21} &\leq \left(\sum_{k=0}^{m-1}\left(\int_{t_k}^{t_{k+1}}\left(\mathbf{E}\|S_{h,\Delta t}^{(m-k)}P_hF'(X(t_k))(O(s)-S(s-t_k)O(t_k))\|^2\right)^{1/2}ds\right)^2\right)^{1/2}\\
&\leq C\left(\sum_{k=0}^{m-1}\left(\int_{t_k}^{t_{k+1}}\left(\mathbf{E}\|P_hF'(X(t_k))(O(s)-S(s-t_k)O(t_k))\|^2\right)^{1/2}ds\right)^2\right)^{1/2}.
\end{aligned}
$$

Using Hölder's inequality, the following inequality holds

$$
\left(\int_a^b f(x)dx\right)^2 \leq (b-a)\int_a^b f(x)^2dx, \tag{3.13}
$$

by assuming that f and f^2 are integrable in the bounded interval $[a,b]$. Using (3.13) with Assumption 2.3(a) and Proposition 2.1 yields

$$
\begin{aligned}
I_3^{21} &\leq C\Delta t^{1/2}\left(\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\mathbf{E}\|P_hF'(X(t_k))(O(s)-S(s-t_k)O(t_k))\|^2ds\right)^{1/2}\\
&\leq C\Delta t^{1/2}\left(\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\mathbf{E}\|(O(s)-S(s-t_k)O(t_k))\|^2ds\right)^{1/2}\\
&\leq C\Delta t^{1/2}\left(\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\left(\left(\mathbf{E}\|O(s)-O(t_k)\|^2\right)^{1/2}+\left(\mathbf{E}\|(S(s-t_k)-\mathbf{I})O(t_k))\|^2\right)^{1/2}\right)^2ds\right)^{1/2}.
\end{aligned}
$$

Using Assumption 2.5(b) yields

$$
\begin{aligned}
I_3^{21} &\leq C\Delta t^{1/2}\left(\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\left((s-t_k)^{\theta}+(s-t_k)^{r/2}\left(\mathbf{E}\|O(t_k)\|_r^2\right)^{1/2}\right)^2ds\right)^{1/2}\\
&\leq C\Delta t^{1/2+\theta}.
\end{aligned}
$$

Let us estimate I_3^{22}.

$$
\begin{aligned}
I_3^{22} &\leq \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\left(\mathbf{E}\|S_{h,\Delta t}^{(m-k)}P_h(-A)^{1/2}(-A)^{-1/2}F'(X(t_k))(S(s-t_k)-\mathbf{I})O(t_k))\|^2\right)^{1/2}ds\\
&\leq \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\|S_{h,\Delta t}^{(m-k)}P_h(-A)^{1/2}\|_{L(L^2(\Omega))}\left(\mathbf{E}\|(-A)^{-1/2}F'(X(t_k))(S(s-t_k)-\mathbf{I})O(t_k))\|^2\right)^{1/2}ds.
\end{aligned}
$$

Since $P_h(-A)^{1/2}=(-A_h)^{1/2}$ and $S_{h,\Delta t}$ satisfies the smoothing properties analogous to $S(t)$ independently of h (see for example [16, 6]), and in particular

$$
\|S_{h,\Delta t}^m(-A_h)^{1/2}\|_{L(L^2(\Omega))}=\|(-A_h)^{1/2}S_{h,\Delta t}^m\|_{L(L^2(\Omega))}\leq Ct_m^{-1/2},\quad t_m=m\Delta t>0,
$$

we therefore have

$$
I_3^{22}\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}(t_m-t_k)^{-1/2}\left(\mathbf{E}\|(-A)^{-1/2}F'(X(t_k))((S(s-t_k)-\mathbf{I})O(t_k))\|^2\right)^{1/2}ds.
$$

The usual identification of $H = L^2(\Omega)$ to its dual yields

$$I_3^{22} \leq C \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (t_m - t_k)^{-1/2} \times \left(\mathbf{E} \left(\sup_{\|v\| \leq 1} |\langle v, (-A)^{-1/2} F'(X(t_k))((S(s-t_k) - \mathbf{I})O(t_k))\rangle| \right)^2 \right)^{1/2} ds,$$

where $\langle , \rangle = (,)$ and we change the notation merely to emphasize that H is identified to its dual space. The fact that $(-A)^{-1/2}$ is self-adjoint implies that $\left((-A)^{-1/2} F'(X)\right)^* = F'(X)^*(-A)^{-1/2}$. This combined with the fact that $\mathcal{D}((-A)^{1/2}) \subset H$ thus $H = H^* \subset \mathcal{D}((-A)^{-1/2}) = \left(\mathcal{D}((-A)^{1/2})\right)^*$ continuously and Assumption 2.3 yields

$$\begin{aligned} I_3^{22} &\leq C \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (t_m - t_k)^{-1/2} \\ &\quad \times \left(\mathbf{E} \left(\sup_{\|v\| \leq 1} |\langle F'(X(t_k))^*(-A)^{-1/2} v, (S(s-t_k) - \mathbf{I})O(t_k)\rangle| \right)^2 \right)^{1/2} ds \\ &\leq C \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (t_m - t_k)^{-1/2} \\ &\quad \times \left(\mathbf{E} \left(\sup_{\|v\| \leq 1} \|F'(X(t_k))^*(-A)^{-1/2} v\|_1 \|(S(s-t_k) - \mathbf{I})O(t_k))\|_{-1} \right)^2 \right)^{1/2} ds. \end{aligned}$$

We also have

$$\begin{aligned} I_3^{22} &\leq C \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (t_m - t_k)^{-1/2} \left(\mathbf{E} \left(1 + \|X(t_k)\|_1\right)^2 \| \left(S(s-t_k) - \mathbf{I})O(t_k)\|_{-1}\right)^2 \right)^{1/2} ds \\ &\leq C \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (t_m - t_k)^{-1/2} \left(\mathbf{E} \left(1 + \|X(t_k)\|_1\right)^4 \right)^{1/4} \left(\mathbf{E} \left(\|S(s-t_k) - \mathbf{I})O(t_k)\|_{-1}\right)^4 \right)^{1/4} ds. \\ &\leq C \sum_{k=0}^{m-1} (t_m - t_k)^{-1/2} \left(1 + \left(\mathbf{E} \|X(t_k)\|_1^4 \right)^{1/4} \right) \int_{t_k}^{t_{k+1}} \left(\mathbf{E} \left(\|S(s-t_k) - \mathbf{I})O(t_k)\|_{-1}\right)^4 \right)^{1/4} ds \\ &\leq C \sum_{k=0}^{m-1} (t_m - t_k)^{-1/2} \int_{t_k}^{t_{k+1}} \|(-A)^{-(r/2+1/2)} \left(S(s-t_k) - \mathbf{I}\right) \|_{L(L^2(\Omega))} \left(\mathbf{E} \|O(t_k)\|_r^4 \right)^{1/4} ds \\ &\leq C \sum_{k=0}^{m-1} (t_m - t_k)^{-1/2} \int_{t_k}^{t_{k+1}} \|(-A)^{1/2 - r/2} (-A)^{-1} (S(s-t_k) - \mathbf{I})\|_{L(L^2(\Omega))} ds. \end{aligned}$$

Using Proposition 2.1 and the fact that $(-A)^{1/2-r/2}$ is bounded as $r \in \{1, 2\}$ yields

$$I_3^{22} \leq C \sum_{k=0}^{m-1} (t_m - t_k)^{-1/2} \int_{t_k}^{t_{k+1}} (s - t_k)\, ds = C \Delta t^{3/2} \sum_{k=0}^{m-1} (m-k)^{-1/2}.$$

As the sum above can be bounded by $2M^{1/2}$ we have

$$I_3^{21}+I_3^{22}\leq C(\Delta t+\Delta t^{1/2+\theta})\leq C(\Delta t^{2\theta}).$$

Let us estimate I_3^{23}. Using the fact that $S_{h,\Delta t}^{(m-k)}$ is bounded for any m,k with Assumption 2.3 and Assumption 2.5 yields (with $G=P_hF''(X(t_k)+r(O(s)-O(t_k)))(O(s)-O(t_k),O(s)-O(t_k))$)

$$\begin{aligned}
I_3^{23} &\leq \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\|S_{h,\Delta t}^{(m-k)}\|_{L(L^2(\Omega))}\int_0^1\left(\mathbf{E}\|G\|^2\right)^{1/2}drds\\
&\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\int_0^1\left(\mathbf{E}\|O(s)-O(t_k)\|_{\mathcal{V}}^4\right)^{1/2}drds\\
&\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\left(\left(\mathbf{E}\|O(s)-O(t_k)\|_{\mathcal{V}}^4\right)^{1/4}\right)^2ds\\
&\leq \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}(s-t_k)^{2\theta}\,ds\leq C(\Delta t)^{2\theta}.
\end{aligned}$$

Combining $I_3^{21}+I_3^{22}$ and I_3^{23} yields the following estimate

$$\mathbf{E}\left(\|I_3\|^2\right)^{1/2}\leq C(\Delta t^{2\theta}).$$

We now estimate I_4. Using equation (3.6) of Lemma 3.1, if $F(X(t))\in L_2(\mathbb{D},\mathbb{H})$ with $\sup\limits_{0\leq s\leq T}\mathbf{E}\|F(X(s))\|_1^2<\infty$ (Assumption 2.7), we have

$$\begin{aligned}
\left(\mathbf{E}\|I_4\|^2\right)^{1/2} &\leq \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\left(\mathbf{E}\|T_{m-k}F(X(s))\|_1^2\right)^{1/2}ds\\
&\leq C(h^2\sup_{0\leq s\leq T}\left(\mathbf{E}\|F(X(s))\|_1^2\right)^{1/2}\left(\Delta t\sum_{k=0}^{m-1}t_{m-k}^{-1/2}\right).
\end{aligned}$$

Note that we can bound $\Delta t\sum_{k=0}^{m-1}t_{m-k}^{-1/2}$ by $2\sqrt{T}$ then

$$\left(\mathbf{E}\|I_4\|^2\right)^{1/2}\leq C(h^2+\Delta t)\left(\sup_{0\leq s\leq T}\mathbf{E}\|F(X(s))\|_1^2\right)^{1/2}.$$

Finally we estimate $\left(\mathbf{E}\|I_5\|^2\right)^{1/2}$. Using Proposition 2.1, we have for $0\leq t_1<t_2\leq T$

$$\|S(t_2)-S(t_1)\|_{L(L^2(\Omega))}=\|(-A)S(t_1)(-A)^{-1}\left(\mathbf{I}-S(t_2-t_1)\right)\|_{L(L^2(\Omega))}\leq\frac{(t_2-t_1)}{t_1},$$

then

$$
\begin{aligned}
\left(\mathbf{E}\|I_5\|^2\right)^{1/2} &\leq \left(\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} \|S(t_m-s)-S(t_m-t_k)\|_{L(L^2(\Omega))}ds\right)\left(\sup_{0\leq s\leq T}\mathbf{E}\|F(X(s))\|^2\right)^{1/2} \\
&\leq C\left(\Delta t+\sum_{k=0}^{m-2}\int_{t_k}^{t_{k+1}}\left(\frac{s-t_k}{t_m-s}\right)ds\right) \\
&\leq C\left(\Delta t+\sum_{k=0}^{m-2}\left((m-k-1)\Delta t\right)^{-1}\int_{t_k}^{t_{k+1}}(s-t_k)\,ds\right) \\
&\leq C\left(\Delta t+\Delta t\sum_{k=0}^{m-2}(m-k-1)^{-1}\right).
\end{aligned}
$$

Noting that the sum above is bounded by $\ln(M)$ we have

$$
\left(\mathbf{E}\|I_5\|^2\right)^{1/2} \leq C(\Delta t+\Delta t|\ln(\Delta t)|).
$$

For $F(X(t)) \in L_2(\mathbb{D},\mathbb{H})$ with $\sup\limits_{0\leq s\leq T}\mathbf{E}\|F(X(s)\|_1^2<\infty$ (Assumption 2.7), combining the previous estimates for the term I yields: If $X_0\in L_2(\mathbb{D},\mathcal{D}(-A))$,

$$
\left(\mathbf{E}\|I\|^2\right)^{1/2} \leq C\left(h^2+\Delta t^{2\theta}+\Delta t|\ln(\Delta t)|+\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\left(\mathbf{E}\|X(t_k)-X_k^h\|^2\right)^{1/2}ds\right).
$$

If $X_0\in L_2(\mathbb{D},\mathcal{D}((-A)^\gamma))$

$$
\left(\mathbf{E}\|I\|^2\right)^{1/2} \leq C\left(t_m^{-1/2}(h^2+\Delta t^{\sigma})+\Delta t|\ln(\Delta t)|+\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}}\left(\mathbf{E}\|X(t_k)-X_k^h\|^2\right)^{1/2}ds\right),
$$

where $\sigma=\min(2\theta,\gamma)$.

Finally we combine all our estimates on I, II and III to get $\left(\mathbf{E}\|I\|^2\right)^{1/2}$, $\left(\mathbf{E}\|II\|^2\right)^{1/2}$ and $\left(\mathbf{E}\|III\|^2\right)^{1/2}$ and use the discrete Gronwall lemma to complete the proof.

3.3. Proof of Theorem 2.10. We now prove convergence in $H^1(\Omega)$ and estimate $\left(\mathbf{E}\|X(t_m)-X_m^h\|^2\right)^{1/2}$. For the proof we follow the same steps as in previous section for Theorem 2.9. We now estimate (3.11) in the H^1 norm.

The estimates of the terms II and III follow as in Section 3.2 and we find

$$
(\mathbf{E}\|II\|^2_{H^1(\Omega)})^{1/2}\leq Ch \qquad \text{and} \qquad \left(\mathbf{E}\|III\|^2_{H^1(\Omega)}\right)^{1/2}\leq C\left(\inf_{j\in\mathbb{N}^d\setminus\mathcal{I}_N}\lambda_j\right)^{-1/2}.
$$

We concentrate instead on estimating the first term $I=I_1+I_2+I_3+I_4+I_5$ in (3.12) and estimates on I_1 follow immediately from Lemma 3.1.

If F satisfies Assumption 2.3 (b), then using the Lipschitz condition, the triangle inequality, the fact that P_h is an bounded operator and $S_{h,\Delta t}$ satisfies the smoothing property analogous to $S(t)$ independently of h [16], ie

$$
\|S^m_{h,\Delta t}v\|^2_{H^1(\Omega)}\leq Ct_m^{-1/2}\|v\| \qquad v\in V_h \qquad t_m>0,
$$

we have

$$
\begin{aligned}
(\mathbf{E}\|I_2\|^2_{H^1(\Omega)})^{1/2} &\leq \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} (\mathbf{E}\|S_{h,\Delta t}^{(m-k)}P_h(F(X(t_k)) - F((Z_k^h + P_hP_NO(t_k))))\|^2_{H^1(\Omega)})^{1/2}ds\\
&\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} t_{m-k}^{-1/2}(\mathbf{E}\|F(X(t_k)) - F((Z_k^h + P_hP_NO(t_k))\|^2)^{1/2}ds\\
&\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} t_{m-k}^{-1/2}(\mathbf{E}\|X(t_k) - X_k^h\|^2_{H^1(\Omega)})^{1/2}ds.
\end{aligned}
$$

If $X_0 \in L_2(\mathbb{D}, \mathcal{D}(-A))$, again using Lipschitz condition, triangle inequality, smoothing property of $S_{h,\Delta t}$, but with Lemma 3.2 gives

$$
\begin{aligned}
(\mathbf{E}\|I_3\|^2_{H^1(\Omega)})^{1/2} &\leq \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} (\mathbf{E}\|S_{h,\Delta t}^{(m-k)}P_h(F(X(s)) - F(X(t_k))\|^2_{H^1(\Omega)})^{1/2}ds\\
&\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} t_{m-k}^{-1/2}(\mathbf{E}\|F(X(s)) - F(X(t_k))\|)^{1/2}ds\\
&\leq C\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} t_{m-k}^{-1/2}(\mathbf{E}\|X(s) - X(t_k)\|^2_{H^1(\Omega)})^{1/2}ds\\
&\leq C\left(\sum_{k=0}^{m-1} t_{m-k}^{-1/2}\int_{t_k}^{t_{k+1}}(s-t_k)^{(1-2\epsilon)/2}ds\right)\\
&\quad\times\left(\mathbf{E}\|X_0\|_2^2 + \left(\mathbf{E}\sup_{0\leq s\leq T}\|F(X(s))\|_{H^1(\Omega)}\right)^2 + 1\right)\\
&\leq C(\Delta t)^{(1/2-\epsilon)}\left(\Delta t\sum_{k=0}^{m-1} t_{m-k}^{-1/2}\right)\left(\mathbf{E}\|X_0\|_2^2 + \left(\mathbf{E}\sup_{0\leq s\leq T}\|F(X(s))\|_{H^1(\Omega)}\right)^2 + 1\right),
\end{aligned}
$$

$\epsilon \in (0, 1/4)$ small enough.
As in the previous theorem, we use the fact that $\Delta t\sum_{k=0}^{m-1} t_{m-k}^{-1/2} \leq 2\sqrt{T}$.
For $F(X(t)) \in L_2(\mathbb{D}, \mathbb{H})$ with $\sup_{0\leq s\leq T}\mathbf{E}\|F(X(s))\|_1^2 < \infty$ (Assumption 2.7), then by relation (3.8) of Lemma 3.1 we find

$$
\begin{aligned}
(\mathbf{E}\|I_4\|^2_{H^1(\Omega)})^{1/2} &\leq \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} (\mathbf{E}\|T_{m-k}F(X(s))\|^2_{H^1(\Omega)})^{1/2}ds\\
&\leq C\Delta t\left(\sum_{k=0}^{m-1} t_{m-k}^{-1/2}h + t_{m-k}^{-1}\Delta t\right)\left(\sup_{0\leq s\leq T}\mathbf{E}\|F(X(s))\|_1^2\right)^{1/2}.
\end{aligned}
$$

Note that $\Delta t\sum_{k=0}^{m-1} t_{m-k}^{-1} \leq \ln(T/\Delta t)$ to get

$$
\begin{aligned}
\left(\mathbf{E}\|I_4\|^2_{H^1(\Omega)}\right)^{1/2} &\leq \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} \left(\mathbf{E}\|T_{m-k}F(X(s))\|_1^2\right)^{1/2}ds\\
&\leq C(h + \Delta t\ln(T/\Delta t))\left(\sup_{0\leq s\leq T}\mathbf{E}\|F(X(s))\|_1^2\right)^{1/2}.
\end{aligned}
$$

Finally, using the equivalency $\|.\|_{H^1(\Omega)} \equiv \|(-A)^{1/2}.\|$ in $\mathcal{D}((-A)^{1/2})$, we obviously have for $0 \leq t_1 < t_2 \leq T$

$$\|S(t_2) - S(t_1)\|_{L(H^1(\Omega))} \leq C\|(-A)^{3/2}S(t_1)(-A)^{-1}\left(\mathbf{I} - S(t_2 - t_1)\right)\|_{L(L^2(\Omega))} \leq C\frac{(t_2 - t_1)}{t_1^{3/2}}$$

so with splitting yields

$$\begin{aligned}
(\mathbf{E}\|I_5\|^2_{H^1(\Omega)})^{1/2} &\leq \left(\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} \|S(t_m - s) - S(t_m - t_k)\|_{L(H^1(\Omega))}ds\right)\left(\sup_{0\leq s\leq T}\mathbf{E}\|F(X(s)\|^2_{H^1(\Omega)}\right)^{1/2}\\
&\leq C\left(\int_{t_{m-1}}^{t_m} \|S(t_m - s) - S(t_m - t_{m-1})\|_{L(H^1(\Omega))}ds\right.\\
&\quad\left.+\sum_{k=0}^{m-2}\int_{t_k}^{t_{k+1}}\left(\frac{s - t_k}{(t_m - s)^{3/2}}\right)ds\right)\\
&\leq C\left(\int_{t_{m-1}}^{t_m}\left((t_m - s)^{-1/2} + \Delta t^{-1/2}\right)ds\right.\\
&\quad\left.+\sum_{k=0}^{m-2}(t_m - t_k - \Delta t)^{-3/2}\int_{t_k}^{t_{k+1}}(s - t_k)ds\right)\\
&\leq C\left(\Delta t^{1/2} + \Delta t^{1/2}\sum_{k=0}^{m-2}(m - k - 1)^{-3/2}\right).
\end{aligned}$$

Since the sum above can be bounded by 2 we have that $\left(\mathbf{E}\|I_5\|^2\right)^{1/2} \leq C\Delta t^{1/2}$.
Combining our estimates, and using that $\Delta t^{(1-\gamma/2)}\ln(T/\Delta t)$ is bounded as $\Delta t \to 0$, for $F(X(t)) \in L_2(\mathbb{D}, \mathbb{H})$ with $\sup_{0\leq s\leq T}\mathbf{E}\|F(X(s))\|_1^2 < \infty$ (Assumption 2.7), we have the following.
If $X_0 \in L_2(\mathbb{D}, \mathcal{D}(-A))$ then

$$(\mathbf{E}\|I\|^2_{H^1(\Omega)})^{1/2} \leq C\left((h + \Delta t^{1/2-\epsilon}t_m^{-1/2}) + \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} t_{m-k}^{-1/2}(\mathbf{E}\|X(t_k) - X_k^h\|^2_{H^1(\Omega)})^{1/2}ds\right).$$

If $X_0 \in L_2(\mathbb{D}, \mathcal{D}(-A))$ with $-AX_0 \in L_2(\mathbb{D}, H^1(\Omega))$ then

$$(\mathbf{E}\|I\|^2_{H^1(\Omega)})^{1/2} \leq C\left((h + \Delta t^{1/2-\epsilon}) + \sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} t_{m-k}^{-1/2}(\mathbf{E}\|X(t_k) - X_k^h\|^2_{H^1(\Omega)})^{1/2}ds\right).$$

where $C > 0$ depending of the T, the initial solution X_0, the mild solution X, the nonlinear function F.
Combining our estimates $\left(\mathbf{E}\|I\|^2\right)^{1/2}$, $\left(\mathbf{E}\|II\|^2\right)^{1/2}$ and $\left(\mathbf{E}\|III\|^2\right)^{1/2}$ and using the discrete Gronwall lemma concludes the proof.

4. Numerical Simulations. We consider for our numerical simulations two examples. Below we consider the error estimates in $L^2(\Omega)$ since this is standard in many of the applications for porous flow and the discrete $L^2(\Omega)$ norm is straightforward to implement for different types of boundary conditions.

4.1. A linear reaction–diffusion equation. As a simple example consider the reaction diffusion equation

$$dX = (D\Delta X - 0.5X)dt + dW \qquad \text{given} \qquad X(0) = X_0,$$

in the time interval $[0, T]$ with diffusion coefficient $D = 1$ and homogeneous Neumann boundary conditions on the domain $\Omega = [0, L_1] \times [0, L_2]$. The eigenfunctions $\{e_i^{(1)} e_j^{(2)}\}_{i,j\geq 0}$ of the operator Δ here are given by

$$e_0^{(l)}(x) = \sqrt{\frac{1}{L_l}}, \qquad e_i^{(l)}(x) = \sqrt{\frac{2}{L_l}} \cos(\lambda_i^{(l)} x), \qquad \lambda_0^{(l)} = 0, \qquad \lambda_i^{(l)} = \frac{i\pi}{L_l}$$

where $l \in \{1, 2\}$, $x \in \Omega$ and $i = 1, 2, 3, \cdots$ with the corresponding eigenvalues $\{\lambda_{i,j}\}_{i,j\geq 0}$ given by $\lambda_{i,j} = (\lambda_i^{(1)})^2 + (\lambda_j^{(2)})^2$. We take $L_1 = L_2 = 1$. Notice that $A = D\Delta$ does not satisfy Assumption 2.2 as 0 is an eigenvalue. During the simulations we need to manage the singularity of (4.2) at $\lambda_0 = 0$ or use the perturbed operator $A = D\Delta + \epsilon \mathbf{I}$, $\epsilon > 0$.
Our function $F(u) = -0.5u$ is linear and obviously satisfies Assumption 2.3 (a) with $\mathcal{V} = L^2(\Omega)$. In our simulation we take $X_0 = 0$, then $X(t) \in L_2(\mathbb{D}, \mathbb{H})$ $t \in [0, T]$ (according to Remark 2.6) and we obviously have $F(X(t)) \in L_2(\mathbb{D}, \mathbb{H})$ $t \in [0, T]$, so that Assumption 2.7 is satisfied. In general we are interested in nonlinear F however for this linear system we can find an exact solution to compare our numerics to.
Recall that we assumed for convenience that the eigenfunctions of covariance operator Q and A are the same. In order to fulful Assumption 2.5 with $\mathcal{V} = L^2(\Omega)$, then $\theta = 1/2$ in Assumption 2.5(b), we take in the representation (1.2)

$$q_{i,j} = \left(i^2 + j^2\right)^{-(r+\delta)}, \qquad r > 0, \text{ and } \delta > 0 \text{ small enough.} \tag{4.1}$$

We obviously have

$$\sum_{(i,j)\in\mathbb{N}^2} \lambda_{i,j}^{r-1} q_{i,j} < \pi^2 \sum_{(i,j)\in\mathbb{N}^2} \left(i^2 + j^2\right)^{-(1+\delta)} < \infty \qquad r \in \{1, 2\},$$

thus Assumption 2.5(a) is satisfied. In the implementation of our modified scheme at every time step, $O(t_{k+1})$ is generated using $O(t_k)$ from the following relation

$$O(t_{k+1}) = e^{A\Delta t} O(t_k) + \int_{t_k}^{t_{k+1}} e^{(t_{k+1}-s)A} dW(s),$$

where $O(0) = 0$. We expand in Fourier space and apply the Ito isometry in each mode and project onto N modes to obtain for $k = 1, 2, \ldots, M-1$

$$(e_i, O(t_{k+1})) = e^{-\lambda_i \Delta t}(e_i, O(t_k)) + \left(\frac{q_i}{2\lambda_i}\left(1 - e^{-2\lambda_i \Delta t}\right)\right)^{1/2} R_{i,k}, \tag{4.2}$$

where $R_{i,k}$ are independent, standard normally distributed random variables with means 0 and variance 1, and $i \in \mathcal{I}_N = \{1, 2, 3, ..., N\}^2$. These are the linear functionals used by Jentzen and Kloeden in [12]. The noise is then projected onto the finite element space by P_h. We evaluated $O(t)$ at the mesh vertices, thus the projection P_h becomes trivial.
In our simulations we examine both the finite element and the finite volume discretization in space. For the cell center finite volume discretization we take $\Delta x = \Delta y =$

1/100. The finite element triangulation was constructed so that the center of the control volume for the finite volume method was a vertex in finite element mesh. In all our simulations in this paper, we denote by 'ModifiedImplicitfem$r$', $r \in \{1, 2\}$ for the graphs with the finite element in space discretization and our modified scheme for time discretization. For the finite volume discretization we have implemented both the new modified method denoted by 'ModifiedImplicitfvm$r$', $r \in \{1, 2\}$ and a standard implicit Euler–Maruyama method denoted 'Implicitfvm$r$', $r \in \{1, 2\}$. In Figure 4.1(a), we see that the observed rate of convergence for the finite element discretization agrees with Theorem 2.9 with $\mathcal{V} = L^2(\Omega)$ and $\theta = 1/2$, and rate of convergence in Δt is very close to 1 for $r \in \{1, 2\}$ in our modified scheme. We also observe that the finite element and finite volume methods for space discretization give the same errors as the mesh is regular. More importantly we see that the error using the new modified scheme is smaller in all the cases than using the standard implicit Euler–Maruyama scheme. Indeed we observe numerically a slower rate of convergence for the the standard scheme of 0.65 for $r = 1$ and 0.98 for $r = 2$ respectively. This is compared to 0.9960 for $r = 1$ and 1.0074 for $r = 2$ with the modified scheme. We also observe that the error decreases as the regularity increases from $r = 1$ to $r = 2$.

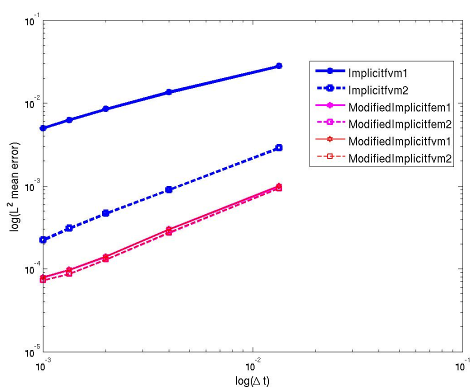


FIG. 4.1. *Convergence in the root mean square L^2 norm at $T = 1$ as a function of Δt. We show convergence for noise where the stochastic process O is in $H^r(\Omega)$ respect to space variable, $r \in \{1, 2\}$ and $\delta = 0.05$ in relation (4.1) for finite element and finite volume space discretizations. We also show convergence of the standard semi–implicit scheme for the finite volume discretization. We used here 30 realizations.*

4.2. Stochastic advection diffusion reaction. As a more challenging example we consider the stochastic advection diffusion reaction SPDE (1.6), with $D = 10^{-2}$ and mixed Neumann-Dirichlet boundary conditions. The Dirichlet boundary condition is $X = 1$ at $\Gamma = \{(x, y) : \ x = 0\}$ and we use the homogeneous Neumann boundary conditions elsewhere. Our goal here is to show that with the well known eigenvalues and eigenfunctions of the operator $\varDelta$ with Neumann (or Dirichlet) boundary conditions, we can apply the new scheme to mixed boundary conditions for the operator $A = D\varDelta$ without explicitly having eigenvalues and eigenfunctions of A. Indeed in term of the abstract setting in (1.1), using the trace operator $\gamma_1 \equiv \dfrac{\partial}{\partial \nu}$ (see [17]) in Green's theorem yields

$$dX = (AX + F_1(X) + b(X))dt + dW, \tag{4.3}$$

where

$$(Au, v) = -\int_{\Omega} D\nabla u \nabla v \, dx, \quad (bu, v) = \int_{\Gamma} \gamma_1 u \gamma_0 v \, d\sigma, \ \gamma_0 v = v \mid_{\partial\Omega}, \ v \in H^1(\Omega), \tag{4.4}$$

$$u \in \{x \in H^2(\Omega) : \frac{\partial x}{\partial \nu} = 0 \ \text{ in } \ \Gamma_1\}, \quad \Gamma_1 = \partial\Omega \diagdown \Gamma. \tag{4.5}$$

In the abstract setting in (1.1), the linear operator is $A = D\Delta$ using only homogeneous Neumann boundary. The explicit expression of b is unknown. The finite volume method (finite element or finite difference methods) uses a natural approximation of b (see [19, 17] for finite volume approximation). The nonlinear term is $F = F_1 + b$ where

$$F_1(u) = -\nabla \cdot (\mathbf{q}u) - \frac{u}{(u+1)}, \quad u \in \mathbb{R}^+. \tag{4.6}$$

As b is linear, F clearly satisfies Assumption 2.3 (b). We use a heterogeneous medium with three parallel high permeability streaks, 100 times higher compared to the other part of the medium. This could represent for example a highly idealized fracture pattern. We obtain the Darcy velocity field $\mathbf{q}$ by solving the system (1.7) with Dirichlet boundary conditions $\Gamma_D^1 = \{0,1\} \times [0,1]$ and Neumann boundary $\Gamma_N^1 = (0,1) \times \{0,1\}$ such that

$$p = \begin{cases} 1 & \text{in} \quad \{0\} \times [0,1] \\ 0 & \text{in} \quad \{L_1\} \times [0,1] \end{cases}$$

$$-k \, \nabla p(\mathbf{x}, t) \cdot \mathbf{n} = 0 \quad \text{in} \quad \Gamma_N^1.$$

To deal with high Péclet flows we discretize in space using finite volumes. We can write the semi-discrete finite volume method as

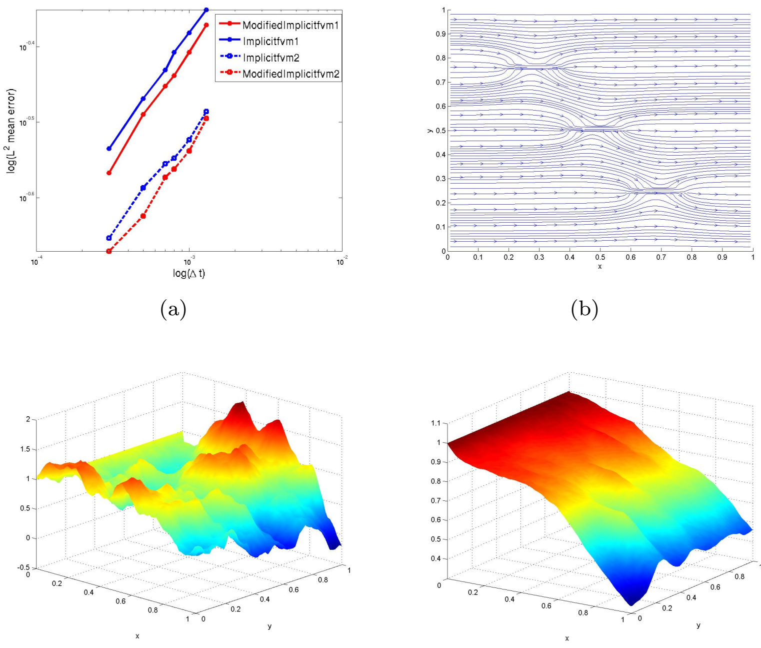
$$dX^h = (A_h X^h + P_h F_1(X^h) + P_h b(X^h)) + P_h P_N dW, \tag{4.7}$$

where here A_h is the space discretization of $D\Delta$ using only homogeneous Neumann boundary conditions and $P_h b(X^h)$ comes from the approximation of diffusion flux on the Dirichlet boundary condition side (see [19, 20]). Thus we can form the noise as in Section 4.1 with the eigenvalues function of Δ with full Neumann boundary conditions and (4.1).

Figure 4.2(a) shows the convergence of the modified method and standard implicit with $O(t) \in L_2(\mathbb{D}, H^r)$, $r \in \{1, 2\}$, $0 \leq t \leq T$. We observe that the temporal convergence order is close to 1/4 for all the schemes. The predicted order 0.5 for $r = 2$ four our modified scheme is probably not achieved due to the fact that Assumption 2.7 is not satisfied. We do note, however, that even if the conditions of the theorem are not fully satisfied we still obtain an improvement in the accuracy over the standard semi-implicit method. Figure 4.2(b) shows the streamline of the velocity field, Figure 4.2(c) shows the mean of 1000 realizations of the "true solution" (with the smallest time step $\Delta t = 1/7680$) for $r = 1$ while Figure 4.2(d) shows a sample of the "true solution" with $r = 1$.

Acknowledgements. We would like to thank Dr A. Jentzen for very useful discussions at an early stage of this paper. These were made possible through an ARC–DAAD grant.

REFERENCES



(c) (d)

Fig. 4.2. *(a) Convergence of the root mean square L^2 norm at $T = 1$ as a function of Δt with 1000 realizations with $\Delta x = \Delta y = 1/100$, $X_0 = 0$. The noise is white in time and the stochastic process $O(t) \in H^r(\Omega)$ respect to space variable, $r \in \{1, 2\}$ and $\delta = 0.05$ in relation (4.1). The temporal order of convergence in time is 1/4, (c) is the streamline of the velocity field. In (b) we plot a sample the 'true solution' for $r = 1$ with $\Delta t = 1/7680$ while (d) is the mean of 1000 realizations.*

[1] E. J. Allen, S. J. Novosel, and Z. Zhang, *Finite element and difference approximation of some linear stochastic partial differential equations*, Stochastics Stochastics Rep., 64 (1998), pp. 117–142.

[2] J. Bear, *Dynamics of Fluids in Porous Media*, Dover, 1988.

[3] P.B. Bedient, H.S. Rifai, and C.J. Newell, *Ground Water Contamination: Transport and Remediation*, Prentice Hall PTR , Englewood Cliffs, New Jersey 07632, 1994.

[4] P-L. Chow, *Stochastic Partial Differential Equations*, Applied Mathematics and nonlinear Science, Chapman & Hall / CRC, 2007. ISBN-1-58488-443-6.

[5] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, vol. 44 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1992.

[6] C. M. Elliott and S. Larsson, *Error estimates with smooth and nonsmooth data for a finite element method for the cahn-hilliard equation*, Math. Comp., (1992).

[7] H. Fujita and T. Suzuki, *Evolutions problems (part1)*, in Handbook of Numerical Analysis, J. L. Lions P. G. Ciarlet, ed., vol. II, North-Holland, 1991, pp. 789–928.

[8] E. Hausenblas, *Approximation for semilinear stochastic evolution equations*, Potential Analysis, 18(2):141–186 (2003).

[9] D. Henry, *Geometric theory of semilinear parabolic equations*, no. 840 in Lecture notes in mathematics, Springer, 1981.

[10] A. Jentzen, *Pathwise numerical approximations of SPDEs*, Potential Analysis, 31(4):375–404 (2009).

[11] ———, *High order pathwise numerical approximations of SPDES with additive noise*, SIAM J. Num. Anal., (2011).

[12] A. Jentzen and P.E. Kloeden, *Overcoming the order barrier in the numerical approximation of SPDEs with additive space-time noise*, Proc. R. Soc. A, 465(2102):649–667 (2009).

[13] A. Jentzen, P. E. Kloeden, and G. Winkel, *Efficient simulation of nonlinear parabolic SPDES with additive noise*, Annals of Applied Probability, To appear (2010).

[14] M. A Katsoulakis, G.T. Kossioris, and O. Lakkis, *Noise regularization and computations for the 1-dimensional stochastic allen–cahn problem*, Interfaces and Free Boundaries, (2007).

[15] G. T. Kossioris and G. E. Zouraris, *Fully-discrete finite element approximations for a fourth-order linear stochastic parabolic equation with additive space-time white noise*, ESAIM, (2010).

[16] S. Larsson, *Nonsmooth data error estimates with applications to the study of the long-time behavior of finite element solutions of semilinear parabolic problems,*. Preprint 1992-36, Department of Mathematics, Chalmers University of Technology, 1992.

[17] P.Knabner and L.Angermann, *Numerical methods for elliptic and parabolic partial differential equations solution*, Springer, 2000.

[18] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Springer, 2007. ISBN-10: 3540707808.

[19] T.Gallouet R.Eymard and R.Herbin, *Finite volume methods*, in: P.G.Ciarlet, J.L.Lions (Eds.)Handbook of Numerical Analysis Volume 7, North-Holland, Amsterdam, pp. 713–1020, 2003.

[20] A. Tambue, *Efficient Numerical Methods for Porous Media Flow*, PhD thesis, Department of Mathematics, Heriot–Watt University, 2010.

[21] V. Thomée, *Galerkin finite element methods for parabolic problems*, Springer Series in Computational Mathematics, 1997.

[22] Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise*, BIT Numerical Mathematics, 44 (2004), pp. 829–847. DOI:10.1007/s10543-004-3755-5.

[23] ———, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM J. Num. Anal., 43 (2005), pp. 1363–1384.